



(12) 发明专利申请

(10) 申请公布号 CN 111898701 A

(43) 申请公布日 2020. 11. 06

(21) 申请号 202010816094.5

(22) 申请日 2020.08.13

(71) 申请人 网易(杭州)网络有限公司

地址 310052 浙江省杭州市滨江区长河街
道网商路599号4幢7层

(72) 发明人 陈伟民 袁焱 范长杰 胡志鹏

(74) 专利代理机构 北京超成律师事务所 11646

代理人 张芮

(51) Int. Cl.

G06K 9/62 (2006.01)

G06K 9/46 (2006.01)

G06N 3/04 (2006.01)

权利要求书3页 说明书18页 附图5页

(54) 发明名称

模型训练、帧图像生成、插帧方法、装置、设备及介质

(57) 摘要

本发明提供一种模型训练、帧图像生成、插帧方法、装置、设备及介质,涉及模型训练技术领域。应用于神经网络模型,采用第一特征提取模块,提取样本前后帧图像的局部特征;采用第二特征提取模块,提取样本前后帧图像的非局部特征;采用帧合成模块,根据局部特征和非局部特征的合成特征,生成样本中间帧图像;根据样本中间帧图像和对应的标签中间帧图像,对神经网络模型进行训练,得到训练后的神经网络模型。基于该方式所训练得到的神经网络模型,扩大了感受野,增强了对前后帧图像中存在较大变化的学习能力,基于该训练好的神经网络模型对前后帧图像处理时,使得生成的中间帧图像更加准确。



1. 一种模型训练方法,其特征在于,应用于神经网络模型,所述神经网络模型包括:第一特征提取模块、多个第二特征提取模块和帧合成模块,包括:

采用各所述第一特征提取模块,提取样本前后帧图像的局部特征;

采用所述第二特征提取模块,提取所述样本前后帧图像的非局部特征;

采用所述帧合成模块,根据所述局部特征和所述非局部特征的合成特征,生成样本中间帧图像;

根据所述样本中间帧图像和对应的标签中间帧图像,对所述神经网络模型进行训练,得到训练后的神经网络模型。

2. 如权利要求1所述的方法,其特征在于,所述第二特征提取模块包括:依次连接的多组第一子模块、自注意力模块、多组第二子模块和上采样层,其中,每组所述第一子模块包括:降采样层和残差模块;每组所述第二子模块包括:所述上采样层和所述残差模块;所述采用所述第二特征提取模块,提取所述样本前后帧图像的非局部特征,包括:

采用多组所述第一子模块,缩小所述样本前后帧图像的尺寸,得到缩小后的特征图;

采用所述自注意力模块,对所述缩小后的特征图进行加权,得到目标特征图,所述目标特征图包含有所述样本前后帧图像的非局部特征;

采用多组所述第二子模块和所述上采样层,放大所述目标特征图。

3. 如权利要求2所述的方法,其特征在于,所述自注意力模块包括:第一卷积层、第二卷积层、第三卷积层、第四卷积层;所述采用所述自注意力模块,对缩小后的所述样本前后帧图像进行加权,得到特征图,包括:

分别采用所述第一卷积层、第二卷积层、第三卷积层对所述缩小后的特征图进行维度调整,得到第一调整结果、第二调整结果、第三调整结果;

根据所述第一调整结果和第二调整结果,确定权重参数;

根据所述权重参数和所述第三调整结果进行加权,得到加权结果;

采用所述第四卷积层对所述加权结果进行维度恢复,得到维度恢复特征,并根据缩小后的所述样本前后帧图像和所述维度恢复特征,确定所述目标特征图。

4. 如权利要求1所述的方法,其特征在于,多个所述第一特征提取模块包括:光流估计模块、深度估计模块、核估计模块;所述采用各所述第一特征提取模块,提取样本前后帧图像的局部特征,包括:

采用所述光流估计模块,计算所述样本前后帧图像的前向光流和后向光流,所述前向光流为样本前帧图像至样本后帧图像的像素位移,所述后向光流为所述样本后帧图像至所述样本前帧图像的像素位移;

采用所述深度估计模块,确定所述样本前后帧图像中对象间的位置关系;

采用所述核估计模块,生成所述样本前后帧图像中各像素的卷积核;所述局部特征包括:所述前向光流、所述后向光流、所述对象间的位置关系、所述各像素的卷积核。

5. 如权利要求4所述的方法,其特征在于,所述神经网络模型包括深度感知投影模块,在所述采用所述帧合成模块,根据所述局部特征和所述非局部特征的合成特征,生成样本中间帧图像之前,所述方法还包括:

采用所述深度感知投影模块,根据所述对象间的位置关系,重新计算所述前向光流和所述后向光流,得到目标前向光流和目标后向光流;

根据所述样本前后帧图像、所述目标前向光流、所述目标后向光流、所述非局部特征、所述各像素的卷积核确定所述合成特征。

6. 如权利要求5所述的方法,其特征在于,所述神经网络模型包括自适应变形层、特征连接层,根据所述样本前后帧图像、所述目标前向光流、所述目标后向光流、所述非局部特征、所述各像素的卷积核确定所述合成特征,包括:

采用所述自适应变形层,根据所述各像素的卷积核、所述前向光流和所述后向光流,对所述样本前后帧图像进行采样,得到采样结果;

采用所述自适应变形层,对所述样本前后帧图像、所述目标前向光流、所述目标后向光流、所述非局部特征进行变形处理,得到多个变形结果;

采用所述特征连接层,对所述采样结果和多个所述变形结果进行连接,得到所述合成特征。

7. 如权利要求1所述的方法,其特征在于,所述帧合成模块包括依次连接的多个残差模块,所述采用所述帧合成模块,根据所述局部特征和所述非局部特征的合成特征,生成样本中间帧图像,包括:

采用多个所述残差模块,根据所述局部特征和所述非局部特征的合成特征,生成所述样本中间帧图像。

8. 如权利要求1所述的方法,其特征在于,根据所述样本中间帧图像和对应的标签中间帧图像,对所述神经网络模型进行训练,包括:

根据所述样本中间帧图像和对应的标签中间帧图像,确定初始神经网络模型的损失函数值;

根据所述损失函数值,对所述神经网络模型的参数进行调节,直至调节后的神经网络模型的损失函数值收敛。

9. 如权利要求8所述的方法,其特征在于,所述根据所述样本中间帧图像和对应的标签中间帧图像,确定初始神经网络模型的损失函数值,包括:

根据所述样本中间帧图像和对应的标签中间帧图像,确定初始神经网络模型的像素损失和对抗损失,所述像素损失用于表示所述样本中间帧图像和对应的标签中间帧图像之间的相似度;

根据所述像素损失和所述对抗损失,确定所述神经网络模型的损失函数值。

10. 如权利要求9所述的方法,其特征在于,所述根据所述样本中间帧图像和对应的标签中间帧图像,确定初始神经网络模型的对抗损失,包括:

采用判别器,确定所述样本中间帧图像比所述标签中间帧图像真实的概率,以及所述标签中间帧图像比所述样本中间帧图像虚假的概率;

根据所述真实的概率和所述虚假的概率,确定所述对抗损失。

11. 一种视频插帧方法,其特征在于,所述方法包括:

从视频中获取目标前后帧图像;

采用预设的神经网络模型对所述目标前后帧图像进行处理,得到目标中间帧图像;其中,所述神经网络模型为采用权利要求1-10任一所述的训练方法得到的所述神经网络模型;

根据所述目标中间帧图像对所述目标前后帧图像进行插帧处理。

12. 根据权利要求11所述的方法,其特征在于,所述从视频中获取目标前后帧图像,包括:

检测所述视频中的至少一个重复帧标识;

根据所述重复帧标识从所述视频中的多个帧图像中,确定所述目标前后帧图像。

13. 如权利要求12所述的方法,其特征在于,所述检测所述视频中的至少一个重复帧标识,包括:

计算所述视频中的相邻两个帧图像之间的像素损失,所述相邻两个帧图像包括:前一帧图像和后一帧图像;

若所述像素损失小于或者等于预设阈值,则将所述后一帧图像的标识确定为所述重复帧标识。

14. 如权利要求13所述的方法,其特征在于,所述根据所述目标中间帧图像对所述目标前后帧图像进行插帧处理,包括:

将所述重复帧标识对应的帧图像,替换为所述目标中间帧图像。

15. 一种模型训练装置,其特征在于,应用于神经网络模型,所述神经网络模型包括:第一特征提取模块、多个第二特征提取模块和帧合成模块,包括:

提取模块,用于采用各所述第一特征提取模块,提取样本前后帧图像的局部特征;采用所述第二特征提取模块,提取所述样本前后帧图像的非局部特征;

生成模块,用于采用所述帧合成模块,根据所述局部特征和所述非局部特征的合成特征,生成样本中间帧图像;

训练模块,用于根据所述样本中间帧图像和对应的标签中间帧图像,对所述神经网络模型进行训练,得到训练后的神经网络模型。

16. 一种视频插帧装置,其特征在于,所述装置包括:

获取模块,用于从视频中获取目标前后帧图像;采用预设的神经网络模型对所述目标前后帧图像进行处理,得到目标中间帧图像;其中,所述神经网络模型为采用权利要求1-10任一所述的训练方法得到的所述神经网络模型;

处理模块,用于根据所述目标中间帧图像对所述目标前后帧图像进行插帧处理。

17. 一种电子设备,其特征在于,包括:存储器和处理器,所述存储器存储有所述处理器可执行的计算机程序,所述处理器执行所述计算机程序时实现上述权利要求1-14任一项所述的方法。

18. 一种存储介质,其特征在于,所述存储介质上存储有计算机程序,所述计算机程序被读取并执行时,实现上述权利要求1-14任一项所述的方法。

模型训练、帧图像生成、插帧方法、装置、设备及介质

技术领域

[0001] 本发明涉及模型训练技术领域,具体而言,涉及一种模型训练、帧图像生成、插帧方法、装置、设备及介质。

背景技术

[0002] 帧速率是指在显示器上显示连续图像的频率,在一定范围内当帧速率越高,画面就越流畅。由于成本问题、硬件限制、网络传输等不可抗力,通常只能获得低帧速率的画面,因此,可以通过插帧的方式在现有帧之间生成中间帧,使得画面更加流畅。

[0003] 相关技术中,通过核估计的方式,对前帧图像和后帧图像中每个像素训练一个卷积核,通过若干独立的卷积核与前后帧进行卷积操作,生成中间帧图像。

[0004] 但是,相关技术中,采用的核估计方式,基于卷积核对图像的局部信息进行处理,当前帧图像和后帧图像变化较大的时,容易造成生成的中间帧图像不准确的问题。

发明内容

[0005] 本发明的目的在于,针对上述现有技术中的不足,提供一种模型训练、帧图像生成、插帧方法、装置、设备及介质,以便解决相关技术中,采用的核估计方式,基于卷积核对图像的局部信息进行处理,当前帧图像和后帧图像变化较大的时,容易造成生成的中间帧图像不准确的问题。

[0006] 为实现上述目的,本发明实施例采用的技术方案如下:

[0007] 第一方面,本发明实施例提供了一种模型训练方法,应用于神经网络模型,所述神经网络模型包括:第一特征提取模块、多个第二特征提取模块和帧合成模块,包括:

[0008] 采用各所述第一特征提取模块,提取样本前后帧图像的局部特征;

[0009] 采用所述第二特征提取模块,提取所述样本前后帧图像的非局部特征;

[0010] 采用所述帧合成模块,根据所述局部特征和所述非局部特征的合成特征,生成样本中间帧图像;

[0011] 根据所述样本中间帧图像和对应的标签中间帧图像,对所述神经网络模型进行训练,得到训练后的神经网络模型。

[0012] 可选的,所述第二特征提取模块包括:依次连接的多组第一子模块、自注意力模块、多组第二子模块和上采样层,其中,每组所述第一子模块包括:降采样层和残差模块;每组所述第二子模块包括:所述上采样层和所述残差模块;所述采用所述第二特征提取模块,提取所述样本前后帧图像的非局部特征,包括:

[0013] 采用多组所述第一子模块,缩小所述样本前后帧图像的尺寸,得到缩小后的特征图;

[0014] 采用所述自注意力模块,对所述缩小后的特征图进行加权,得到目标特征图,所述目标特征图包含有所述样本前后帧图像的非局部特征;

[0015] 采用多组所述第二子模块和所述上采样层,放大所述目标特征图。

[0016] 可选的,所述自注意力模块包括:第一卷积层、第二卷积层、第三卷积层、第四卷积层;所述采用所述自注意力模块,对缩小后的所述样本前后帧图像进行加权,得到特征图,包括:

[0017] 分别采用所述第一卷积层、第二卷积层、第三卷积层对所述缩小后的特征图进行维度调整,得到第一调整结果、第二调整结果、第三调整结果;

[0018] 根据所述第一调整结果和第二调整结果,确定权重参数;

[0019] 根据所述权重参数和所述第三调整结果进行加权,得到加权结果;

[0020] 采用所述第四卷积层对所述加权结果进行维度恢复,得到维度恢复特征,并根据缩小后的所述样本前后帧图像和所述维度恢复特征,确定所述目标特征图。

[0021] 可选的,多个所述第一特征提取模块包括:光流估计模块、深度估计模块、核估计模块;所述采用各所述第一特征提取模块,提取样本前后帧图像的局部特征,包括:

[0022] 采用所述光流估计模块,计算所述样本前后帧图像的前向光流和后向光流,所述前向光流为样本前帧图像至样本后帧图像的像素位移,所述后向光流为所述样本后帧图像至所述样本前帧图像的像素位移;

[0023] 采用所述深度估计模块,确定所述样本前后帧图像中对象间的位置关系;

[0024] 采用所述核估计模块,生成所述样本前后帧图像中各像素的卷积核;所述局部特征包括:所述前向光流、所述后向光流、所述对象间的位置关系、所述各像素的卷积核。

[0025] 可选的,所述神经网络模型包括深度感知投影模块,在所述采用所述帧合成模块,根据所述局部特征和所述非局部特征的合成特征,生成样本中间帧图像之前,所述方法还包括:

[0026] 采用所述深度感知投影模块,根据所述对象间的位置关系,重新计算所述前向光流和所述后向光流,得到目标前向光流和目标后向光流;

[0027] 根据所述样本前后帧图像、所述目标前向光流、所述目标后向光流、所述非局部特征、所述各像素的卷积核确定所述合成特征。

[0028] 可选的,所述神经网络模型包括自适应变形层、特征连接层,根据所述样本前后帧图像、所述目标前向光流、所述目标后向光流、所述非局部特征、所述各像素的卷积核确定所述合成特征,包括:

[0029] 采用所述自适应变形层,根据所述各像素的卷积核、所述前向光流和所述后向光流,对所述样本前后帧图像进行采样,得到采样结果;

[0030] 采用所述自适应变形层,对所述样本前后帧图像、所述目标前向光流、所述目标后向光流、所述非局部特征进行变形处理,得到多个变形结果;

[0031] 采用所述特征连接层,对所述采样结果和多个所述变形结果进行连接,得到所述合成特征。

[0032] 可选的,所述帧合成模块包括依次连接的多个残差模块,所述采用所述帧合成模块,根据所述局部特征和所述非局部特征的合成特征,生成样本中间帧图像,包括:

[0033] 采用多个所述残差模块,根据所述局部特征和所述非局部特征的合成特征,生成所述样本中间帧图像。

[0034] 可选的,根据所述样本中间帧图像和对应的标签中间帧图像,对所述神经网络模型进行训练,包括:

- [0035] 根据所述样本中间帧图像和对应的标签中间帧图像,确定初始神经网络模型的损失函数值;
- [0036] 根据所述损失函数值,对所述神经网络模型的参数进行调节,直至调节后的神经网络模型的损失函数值收敛。
- [0037] 可选的,所述根据所述样本中间帧图像和对应的标签中间帧图像,确定初始神经网络模型的损失函数值,包括:
- [0038] 根据所述样本中间帧图像和对应的标签中间帧图像,确定初始神经网络模型的像素损失和对抗损失,所述像素损失用于表示所述样本中间帧图像和对应的标签中间帧图像之间的相似度;
- [0039] 根据所述像素损失和所述对抗损失,确定所述神经网络模型的损失函数值。
- [0040] 可选的,所述根据所述样本中间帧图像和对应的标签中间帧图像,确定初始神经网络模型的对抗损失,包括:
- [0041] 采用判别器,确定所述样本中间帧图像比所述标签中间帧图像真实的概率,以及所述标签中间帧图像比所述样本中间帧图像虚假的概率;
- [0042] 根据所述真实的概率和所述虚假的概率,确定所述对抗损失。
- [0043] 第二方面,本发明实施例还提供了一种视频插帧方法,所述方法包括:
- [0044] 从视频中获取目标前后帧图像;
- [0045] 采用预设的神经网络模型对所述目标前后帧图像进行处理,得到目标中间帧图像;其中,所述神经网络模型为采用第一方面任一所述的训练方法得到的所述神经网络模型;
- [0046] 根据所述目标中间帧图像对所述目标前后帧图像进行插帧处理。
- [0047] 可选的,所述从视频中获取目标前后帧图像,包括:
- [0048] 检测所述视频中的至少一个重复帧标识;
- [0049] 根据所述重复帧标识从所述视频中的多个帧图像中,确定所述目标前后帧图像。
- [0050] 可选的,所述检测所述视频中的至少一个重复帧标识,包括:
- [0051] 计算所述视频中的相邻两个帧图像之间的像素损失,所述相邻两个帧图像包括:前一帧图像和后一帧图像;
- [0052] 若所述像素损失小于或者等于预设阈值,则将所述后一帧图像的标识确定为所述重复帧标识。
- [0053] 可选的,所述根据所述目标中间帧图像对所述目标前后帧图像进行插帧处理,包括:
- [0054] 将所述重复帧标识对应的帧图像,替换为所述目标中间帧图像。
- [0055] 第三方面,本发明实施例还提供了一种模型训练装置,应用于神经网络模型,所述神经网络模型包括:第一特征提取模块、多个第二特征提取模块和帧合成模块,包括:
- [0056] 提取模块,用于采用各所述第一特征提取模块,提取样本前后帧图像的局部特征;采用所述第二特征提取模块,提取所述样本前后帧图像的非局部特征;
- [0057] 生成模块,用于采用所述帧合成模块,根据所述局部特征和所述非局部特征的合成特征,生成样本中间帧图像;
- [0058] 训练模块,用于根据所述样本中间帧图像和对应的标签中间帧图像,对所述神经

网络模型进行训练,得到训练后的神经网络模型。

[0059] 可选的,所述第二特征提取模块包括:依次连接的多组第一子模块、自注意力模块、多组第二子模块和上采样层,其中,每组所述第一子模块包括:降采样层和残差模块;每组所述第二子模块包括:所述上采样层和所述残差模块;所述提取模块,还用于采用多组所述第一子模块,缩小所述样本前后帧图像的尺寸,得到缩小后的特征图;采用所述自注意力模块,对所述缩小后的特征图进行加权,得到目标特征图,所述目标特征图包含有所述样本前后帧图像的非局部特征;采用多组所述第二子模块和所述上采样层,放大所述目标特征图。

[0060] 可选的,所述自注意力模块包括:第一卷积层、第二卷积层、第三卷积层、第四卷积层;所述提取模块,还用于分别采用所述第一卷积层、第二卷积层、第三卷积层对所述缩小后的特征图进行维度调整,得到第一调整结果、第二调整结果、第三调整结果;根据所述第一调整结果和第二调整结果,确定权重参数;根据所述权重参数和所述第三调整结果进行加权,得到加权结果;采用所述第四卷积层对所述加权结果进行维度恢复,得到维度恢复特征,并根据缩小后的所述样本前后帧图像和所述维度恢复特征,确定所述目标特征图。

[0061] 可选的,多个所述第一特征提取模块包括:光流估计模块、深度估计模块、核估计模块;所述提取模块,还用于采用所述光流估计模块,计算所述样本前后帧图像的前向光流和后向光流,所述前向光流为样本前帧图像至样本后帧图像的像素位移,所述后向光流为所述样本后帧图像至所述样本前帧图像的像素位移;采用所述深度估计模块,确定所述样本前后帧图像中对象间的位置关系;采用所述核估计模块,生成所述样本前后帧图像中各像素的卷积核;所述局部特征包括:所述前向光流、所述后向光流、所述对象间的位置关系、所述各像素的卷积核。

[0062] 可选的,所述神经网络模型包括深度感知投影模块,在所述采用所述帧合成模块,所述装置还包括:

[0063] 计算模块,用于采用所述深度感知投影模块,根据所述对象间的位置关系,重新计算所述前向光流和所述后向光流,得到目标前向光流和目标后向光流;

[0064] 确定模块,用于根据所述样本前后帧图像、所述目标前向光流、所述目标后向光流、所述非局部特征、所述各像素的卷积核确定所述合成特征。

[0065] 可选的,所述神经网络模型包括自适应变形层、特征连接层,所述确定模块,还用于采用所述自适应变形层,根据所述各像素的卷积核、所述前向光流和所述后向光流,对所述样本前后帧图像进行采样,得到采样结果;采用所述自适应变形层,对所述样本前后帧图像、所述目标前向光流、所述目标后向光流、所述非局部特征进行变形处理,得到多个变形结果;采用所述特征连接层,对所述采样结果和多个所述变形结果进行连接,得到所述合成特征。

[0066] 可选的,所述帧合成模块包括依次连接的多个残差模块,所述生成模块,还用于采用多个所述残差模块,根据所述局部特征和所述非局部特征的合成特征,生成所述样本中间帧图像。

[0067] 可选的,所述训练模块,还用于根据所述样本中间帧图像和对应的标签中间帧图像,确定初始神经网络模型的损失函数值;根据所述损失函数值,对所述神经网络模型的参数进行调节,直至调节后的神经网络模型的损失函数值收敛。

[0068] 可选的,所述训练模块,还用于根据所述样本中间帧图像和对应的标签中间帧图像,确定初始神经网络模型的像素损失和对抗损失,所述像素损失用于表示所述样本中间帧图像和对应的标签中间帧图像之间的相似度;

[0069] 根据所述像素损失和所述对抗损失,确定所述神经网络模型的损失函数值。

[0070] 可选的,所述训练模块,还用于采用判别器,确定所述样本中间帧图像比所述标签中间帧图像真实的概率,以及所述标签中间帧图像比所述样本中间帧图像虚假的概率;根据所述真实的概率和所述虚假的概率,确定所述对抗损失。

[0071] 第四方面,本发明实施例还提供了一种视频插帧装置,所述装置包括:

[0072] 获取模块,用于从视频中获取目标前后帧图像;采用预设的神经网络模型对所述目标前后帧图像进行处理,得到目标中间帧图像;其中,所述神经网络模型为采用第一方面任一所述的训练方法得到的所述神经网络模型;

[0073] 处理模块,用于根据所述目标中间帧图像对所述目标前后帧图像进行插帧处理。

[0074] 可选的,所述获取模块,还用于检测所述视频中的至少一个重复帧标识;

[0075] 根据所述重复帧标识从所述视频中的多个帧图像中,确定所述目标前后帧图像。

[0076] 可选的,所述获取模块,还用于计算所述视频中的相邻两个帧图像之间的像素损失,所述相邻两个帧图像包括:前一帧图像和后一帧图像;若所述像素损失小于或者等于预设阈值,则将所述后一帧图像的标识确定为所述重复帧标识。

[0077] 可选的,所述处理模块,还用于将所述重复帧标识对应的帧图像,替换为所述目标中间帧图像。

[0078] 第五方面,本发明实施例还提供了一种电子设备,包括:存储器和处理器,所述存储器存储有所述处理器可执行的计算机程序,所述处理器执行所述计算机程序时实现上述第一方面和第二方面任一项所述的方法。

[0079] 第六方面,本发明实施例还提供了一种存储介质,所述存储介质上存储有计算机程序,所述计算机程序被读取并执行时,实现上述第一方面和第二方面任一项所述的方法。

[0080] 本发明的有益效果是:本申请实施例提供一种模型训练方法,应用于神经网络模型,采用第一特征提取模块,提取样本前后帧图像的局部特征;采用第二特征提取模块,提取样本前后帧图像的非局部特征;采用帧合成模块,根据局部特征和非局部特征的合成特征,生成样本中间帧图像;根据样本中间帧图像和对应的标签中间帧图像,对神经网络模型进行训练,得到训练后的神经网络模型。在提取样本前后帧图像的局部特征的基础上,还通过第二特征提取模块提取非局部特征,基于该方式所训练得到的神经网络模型,扩大了感受野,增强了对前后帧图像中存在较大变化的学习能力,基于该训练好的神经网络模型对变化较大的前后帧图像进行处理时,可以使得生成的中间帧图像更加准确。

附图说明

[0081] 为了更清楚地说明本发明实施例的技术方案,下面将对实施例中所需要使用的附图作简单地介绍,应当理解,以下附图仅示出了本发明的某些实施例,因此不应被看作是对范围的限定,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他相关的附图。

[0082] 图1为本发明实施例提供的一种模型训练方法的流程示意图;

- [0083] 图2为本发明实施例提供的第二特征提取模块的结构示意图；
- [0084] 图3为本发明实施例提供的一种模型训练方法的流程示意图；
- [0085] 图4为本发明实施例提供的自注意力模块的处理示意图；
- [0086] 图5为本发明实施例提供的一种模型训练方法的流程示意图；
- [0087] 图6为本发明实施例提供的神经网络模型的生成器的结构示意图；
- [0088] 图7为本发明实施例提供的一种模型训练方法的流程示意图；
- [0089] 图8为本发明实施例提供的一种模型训练方法的流程示意图；
- [0090] 图9为本发明实施例提供的一种模型训练方法的流程示意图；
- [0091] 图10为本发明实施例提供的一种模型训练方法的流程示意图；
- [0092] 图11为本发明实施例提供的一种模型训练方法的流程示意图；
- [0093] 图12为本发明实施例提供的一种模型训练方法的流程示意图；
- [0094] 图13为本发明实施例提供的一种视频插帧方法的流程示意图；
- [0095] 图14为本发明实施例提供的一种模型训练装置的结构示意图；
- [0096] 图15为本发明实施例提供的一种视频插帧装置的结构示意图；
- [0097] 图16为本发明实施例提供的一种电子设备的结构示意图。

具体实施方式

[0098] 为使本发明实施例的目的、技术方案和优点更加清楚，下面将结合本发明实施例中的附图，对本发明实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例是本发明一部分实施例，而不是全部的实施例。

[0099] 因此，以下对在附图中提供的本申请的实施例的详细描述并非旨在限制要求保护的本申请的范围，而是仅仅表示本申请的选定实施例。基于本申请中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

[0100] 需要说明的是，本发明的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象，而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换，以便这里描述的本发明的实施例能够以除了在这里图示或描述的那些以外的顺序实施。此外，术语“包括”和“具有”以及他们的任何变形，意图在于覆盖不排他的包含，例如，包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元，而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

[0101] 此外，在不冲突的情况下，本申请的实施例中的特征可以相互结合。

[0102] 帧速率是指在显示器上显示连续图像的频率，在一定范围内当帧速率越高，画面就越流畅。由于成本问题、硬件限制、网络传输等不可抗力，通常只能获得低帧速率的画面，因此，可以通过插帧的方式在现有帧之间生成中间帧，使得视频的画面更加流畅。

[0103] 相关技术中，通过核估计的方式，对前帧图像和后帧图像中每个像素训练一个卷积核，通过若干独立的卷积核与前后帧进行卷积操作，生成中间帧图像。但是，相关技术中，采用的核估计方式，基于卷积核对图像的局部信息进行处理，当前帧图像和后帧图像变化较大的时，容易造成生成的中间帧图像不准确的问题。

[0104] 本申请实施例提供一种模型训练方法,在提取样本前后帧图像的局部特征的基础上,还通过第二特征提取模块提取非局部特征,基于该方式所训练得到的神经网络模型,扩大了感受野,增强了对前后帧图像中存在较大变化的学习能力,使用该训练好的神经网络模型对变化较大的前后帧图像进行处理时,可以使得生成的中间帧图像更加准确,从而解决了相关技术中所存在的问题。

[0105] 下述对本申请实施例所涉及的名词进行解释说明。

[0106] 神经网络:由若干个神经元及其之间的连接(神经网络的参数)组成。卷积神经网络是神经网络的一种,它将若干个神经元组织成一个卷积层,数据从输入开始,靠神经元之间的连接,在若干卷积层中顺序传播,直至最终输出。生成对抗网络具有两个卷积神经网络,分别称为生成器和判别器,它们彼此之间进行博弈,交替迭代优化。

[0107] 视频插帧:一种视频处理技术,即通过插值的方式在现有帧之间生成中间动画帧,使动画更流畅,以弥补运动模糊,生成“假”慢动作效果。该技术可广泛应用于影视后期处理、游戏动画生成以及活动推广宣传中,或者用于清晰化视频中模糊的帧。

[0108] 自注意力机制:自注意力机制是注意力机制的改进,其减少了对外部信息的依赖,更擅长捕捉数据或特征的内部相关性。

[0109] 本申请实施例提供的一种模型训练方法,其执行主体可以为电子设备,该电子设备可以为终端,也可以为服务器,还可以其它类型的处理设备,本申请实施例对此不进行具体限制。当电子设备为终端时,终端可以为台式电脑、笔记本电脑等计算机设备,下述以终端为执行主体,对本申请实施例提供的模型训练方法进行解释说明。

[0110] 图1为本发明实施例提供的一种模型训练方法的流程示意图,应用于神经网络模型,神经网络模型包括:第一特征提取模块、多个第二特征提取模块和帧合成模块,如图1所示,该方法可以包括:

[0111] S101、采用各第一特征提取模块,提取样本前后帧图像的局部特征。

[0112] 其中,样本前后帧图像可以包括:样本前帧图像和样本后帧图像。样本前后帧图像均可以为颜色通道的图像,该颜色通道可以为RGB(Red,Green,Blue,红色,绿色,蓝色)三个通道,样本前帧图像和样本后帧图像可以表示RGB三个通道中各个通道的像素信息。

[0113] 在一些实施方式中,每个第一特征提取模块,所提取的样本前后帧图像的局部特征可以不同。样本前后帧图像的局部特征可以表示样本前后帧图像中各像素和/或各区域的局部变化。

[0114] S102、采用第二特征提取模块,提取样本前后帧图像的非局部特征。

[0115] 其中,该第二特征提取模块中引入了自注意力机制,即该第二特征提取模块为基于自注意力机制的特征提取模块。本申请实施例提供的第二特征提取模块包含一个神经网络。

[0116] 在本申请实施例中,采用第二特征提取模块,提取样本前后帧图像的非局部特征,可以扩大神经网络的感受野,使得在合成中间帧图像时,能够感知更大乃至全局的特征,优化模型对于前后帧图像中可能存在的较大动作变化的能力,提高了神经网络模型对前后帧图像汇总存在较大动作变化的处理能力。

[0117] S103、采用帧合成模块,根据局部特征和非局部特征的合成特征,生成样本中间帧图像。

[0118] 其中,样本中间帧图像也具有RGB三个通道,帧合成模块也可以称为帧合成网络,帧合成模块可以由多个卷积层组成。

[0119] 在一些实施方式中,终端可以对部分局部特征进行处理,得到第一处理结果,对部分局部特征和非局部特征进行处理,得到第二处理结果,继而基于第一处理结果和第二处理结果得到合成特征,继而由多个卷积层组成的帧合成网络可以生成样本中间帧图像。

[0120] S104、根据样本中间帧图像和对应的标签中间帧图像,对神经网络模型进行训练,得到训练后的神经网络模型。

[0121] 其中,标签中间帧图像作为标签,可以引导神经网络模型生成质量接近的中间帧图像。

[0122] 在一种可能的实施中,终端可以根据样本中间帧图像和对应的标签中间帧图像,对第一特征提取模块和第二特征提取模块进行协同训练,对每个第一特征提取模块、第二特征提取模块、帧合成模块中的参数进行优化,实现对神经网络模型的训练,得到训练后的神经网络模型。

[0123] 在实际应用中,终端可以将待处理前后帧图像,即相邻的两帧图像,输入上述训练后的神经网络模型中,该神经网络模型可以对该待处理前后帧图像进行处理,得到中间帧图像。训练后的神经网络模型可以广泛应用于影视后期处理、游戏动画生成以及活动宣传推广中,或者用于清晰化视频中模糊的帧中。

[0124] 需要说明的是,训练后的神经网络模型可以关注待处理前后帧图像的全局信息,当待处理前后帧图像变化较大时,训练后的神经网络模型可以准确的输出中间帧图像。

[0125] 综上所述,本申请实施例提供一种模型训练方法,应用于神经网络模型,采用第一特征提取模块,提取样本前后帧图像的局部特征;采用第二特征提取模块,提取样本前后帧图像的非局部特征;采用帧合成模块,根据局部特征和非局部特征的合成特征,生成样本中间帧图像;根据样本中间帧图像和对应的标签中间帧图像,对神经网络模型进行训练,得到训练后的神经网络模型。在提取样本前后帧图像的局部特征的基础上,还通过第二特征提取模块提取非局部特征,基于该方式所训练得到的神经网络模型,扩大了感受野,增强了对前后帧图像中存在较大变化的学习能力,基于该训练好的神经网络模型对变化较大的前后帧图像进行处理时,可以使得生成的中间帧图像更加准确。

[0126] 可选的,图2为本发明实施例提供的第二特征提取模块的结构示意图,如图2所示,上述第二特征提取模块包括:依次连接的多组第一子模块、自注意力模块205、多组第二子模块和上采样层,其中,每组第一子模块包括:降采样层202和残差模块203;每组第二子模块包括:上采样层204和残差模块203。

[0127] 图3为本发明实施例提供的一种模型训练方法的流程示意图,如图3所示,上述S102的过程,可以包括:

[0128] S301、采用多组第一子模块,缩小样本前后帧图像的尺寸,得到缩小后的特征图。

[0129] 在一些实施方式中,样本前帧图像和样本后帧图像可以通过通道进行连接,样本前帧图像和样本后帧图像各有3通道图像,进行通道连接201后可以得到6通道图像。终端可以采用多组第一子模块,将样本前后帧图像缩小为预设倍数。不同的预设倍数对应的多组第一子模块的结构和/或参数不同。如图2所示,该通道连接201、多组第一子模块、自注意力模块205、多组多组第二子模块和上采样层204依次顺序连接。

[0130] 例如,如图2所示,多组第一子模块的数量可以为4,每组第一子模块可以包括:降采样层202和残差模块203,该降采样层202可以为2倍的平均池化层,即每 2×2 个像素,取4个像素的平均值作为输出(1个像素)则可以将通道中的特征图宽度和高度缩小为原来的二分之一。

[0131] 当然,多组第一子模块还可以为其他结构或者参数,以将通道中的特征图缩小为便于自注意力模块进行处理的其他倍数,本申请实施例对此不进行具体限制。例如,可以缩小为原来的三分之一,或者四分之一。

[0132] S302、采用自注意力模块,对缩小后的特征图进行加权,得到目标特征图,目标特征图包含有样本前后帧图像的非局部特征。

[0133] 其中,自注意力模块可以由多个卷积层构成。自注意力模块通过在空间维度上对整个特征图进行加权,打破了卷积核的局部性。

[0134] 在本申请实施例中,终端可以采用自注意力模块,对缩小后的整个特征图分别在空间维度上进行加权,打破了卷积核的局部性,增加了神经网络的学习能力,以使神经网络可以自适应的关注该缩小后的特征图中的重要部分,得到目标特征图,即非局部特征图。

[0135] S303、采用多组第二子模块和上采样层,放大目标特征图。

[0136] 其中,上采样层包括双线性插值操作与一个卷积层。

[0137] 在一种可能的实施方式中,缩小的倍数和放大倍数是相对应的,放大后的目标特征图应当与通道中所输入的特征图大小相同。根据缩小倍数设定相应的放大倍数,多组第二子模块和上采样层的结构和/或参数也可以根据预设放大倍数进行设定。

[0138] 例如,如图2所示,多组第二子模块的数量可以为3,上采样层的数量可以为1,每组第二子模块可以包括:上采样层和残差模块,目标特征图依次通过3组第二子模块、上采样层,得到目标特征图。上采样层中的双线性插值可以将输入特征图的宽度与高度放大为原来的两倍,并通过上采样层中的卷积层对放大的结果进行微调,继而可以得到放大后的目标特征图。

[0139] 在本申请实施例中,残差模块可以由两个顺序连接的卷积层组成。

[0140] 综上所述,采用多组第一子模块,获取缩小后的特征图,可以便于自注意力模块对整个特征图的加权,采用自注意力模块获取目标特征图,可以进一步扩大感受野,使得合成中间帧图像时能够感知更大乃至全局的特征,以便神经网络模型生成的中间帧图像更加准确。采用多组第二子模块和上采样层,放大目标特征图,可以便于后续对目标特征图的处理,以生成中间帧图像。

[0141] 可选的,图4为本发明实施例提供的自注意力模块的处理示意图,如图4所示,自注意力模块包括:第一卷积层、第二卷积层、第三卷积层、第四卷积层。

[0142] 图5为本发明实施例提供了一种模型训练方法的流程示意图,如图5所示,上述S302的过程可以包括:

[0143] S501、分别采用第一卷积层、第二卷积层、第三卷积层对缩小后的特征图进行维度调整,得到第一调整结果、第二调整结果、第三调整结果。

[0144] 在一些实施方式中,第一卷积层、第二卷积层和第三卷积层的卷积核大小可以相同。终端可以通过第一卷积层、第二卷积层和第三卷积层对输入的缩小后的特征图的维度进行调整,以适应硬件算力,便于后续处理。

[0145] 需要说明的是,第一卷积层、第二卷积层和第三卷积层的卷积核大小可以为预设值,该预设值可以根据经验值进行设定,例如,该预设值可以为 $1*1*1$ 。

[0146] S502、根据第一调整结果和第二调整结果,确定权重参数。

[0147] S503、根据权重参数和第三调整结果进行加权,得到加权结果。

[0148] 其中,终端可以对第一调整结果和第二调整结果进行相乘,得到每个像素的权重参数。

[0149] 在本申请实施例中,终端可以对各像素对应的权重参数和第三调整结果中的各像素相乘,以实现权重参数和所述第三调整结果进行加权,得到加权结果。

[0150] S504、采用第四卷积层对加权结果进行维度恢复,得到维度恢复特征,并根据缩小后的样本前后帧图像和维度恢复特征,确定目标特征图。

[0151] 其中,第四卷积层与上述第一卷积层、第二卷积层和第三卷积层的卷积核大小也可以相同。

[0152] 如图4所示,在本申请实施例中,通过第四卷积层可以将加权结果对应的特征图,恢复至与输入的特征图相同的维度,即恢复至与输入的缩小后的特征图相同的维度,得到维度恢复特征,继而将缩小后的样本前后帧图像和维度恢复特征进行相加,得到目标特征图。

[0153] 综上所述,采用S501至S504的过程,获取目标特征图,打破了卷积核的局部性,使得获取的目标特征图包含非局部特征,整个实现过程也更加高效和准确。

[0154] 可选的,图6为本发明实施例提供的神经网络模型的生成器的结构示意图,如图6所示,生成器可以包括:第一特征提取模块、多个第二特征提取模块,多个第一特征提取模块包括:光流估计模块、深度估计模块、核估计模块。

[0155] 图7为本发明实施例提供的一种模型训练方法的流程示意图,如图7所示,上述S101的过程,可以包括:

[0156] S701、采用光流估计模块,计算样本前后帧图像的前向光流和后向光流。

[0157] 其中,采用光流估计模块可以计算样本前后帧图像中的像素位移信息。像素位移信息包括:前向光流和后向光流。前向光流可以为样本前帧图像至样本后帧图像的像素位移,后向光流可以为样本后帧图像至样本前帧图像的像素位移。

[0158] 另外,光流估计模块的神经网络结构可以为预设神经网络结构,例如,该预设神经网络结构可以为PWC-Net(光流学习网络)结构。

[0159] 需要说明的是,样本前后帧图像的前向光流和后向光流可以为包括有前向光流和后向光流的特征图。

[0160] S702、采用深度估计模块,确定样本前后帧图像中对象间的位置关系。

[0161] 在本申请实施例中,由于单一的光流估计模块无法为生成中间帧的估计提供足够的特征,因此,终端可以采用深度估计模块,估计样本前帧图像和样本后帧图像中对象间的相对关系,以提取样本前后帧图像的深度信息,使得确定的局部特征更加丰富、准确。

[0162] 另外,样本前后帧图像中对象间的位置关系可以为包括有对象间的位置关系的深度图。

[0163] S703、采用核估计模块,生成样本前后帧图像中各像素的卷积核。

[0164] 其中,局部特征包括:前向光流、后向光流、对象间的位置关系、各像素的卷积核。

核估计模块,可以使得光流估计模型的关注区域扩大为一个领域,以增加神经网络模型的学习能力。

[0165] 在一种可能的实施方式中,终端可以采用核估计模块,基于样本前后帧图像中对应位置周围和预设数量个像素,为每个像素训练一个卷积核,以便后续根据各像素的卷积核和样本前后帧图像进行卷积操作,确定中间帧图像。

[0166] 需要说明的是,核估计模块可以使用U-net(一种包含下采样和上采样的网络结构)网络,上述各像素的卷积核可以相同,各像素的卷积核可以为预设值,例如,各像素的卷积核可以为4*4。

[0167] 综上所述,本申请实施例中,采用光流估计模块、深度估计模块、核估计模块对样本前后帧图像进行处理,得到前向光流、后向光流、对象间的位置关系、各像素的卷积核等。使得生成的局部特征包含的信息更多、更全面,基于该局部特征生成中间帧图像时,有利于中间帧图像的准确生成。

[0168] 可选的,在上述S103之前,图8为本发明实施例提供的一种模型训练方法的流程示意图,如图8所示,该方法还可以包括:

[0169] S801、根据对象间的位置关系,重新计算前向光流和后向光流,得到目标前向光流和目标后向光流。

[0170] 其中,神经网络模型的生成器还可以包括深度感知投影模块。

[0171] 在本申请实施例中,由于光流估计模块,计算的前向光流和后向光流不能指向正确的对象或像素,因此,如图6所示,终端可以采用深度感知投影方式,根据对象间的位置关系的深度图,重新计算前向光流和后向光流,以生成指向更准确像素的光流向量,即生成目标前向光流和目标后向光流。

[0172] S802、根据样本前后帧图像、目标前向光流、目标后向光流、非局部特征、各像素的卷积核确定合成特征。

[0173] 需要说明的是,终端可以对样本前后帧图像、目标前向光流和目标后向光流对应的特征图、非局部特征对应的特征图,以及各像素的卷积核确定合成的特征。

[0174] 综上所述,上述合成特征中既包含了详细且准确的局部特征、还包含了非局部的特征,对于变化较大的样本前后帧图像,基于该合成特征确定的样本中间帧图像更加准确。

[0175] 可选的,图9为本发明实施例提供的一种模型训练方法的流程示意图,如图9所示,如图6所示,神经网络模型的生成器可以包括自适应变形层、特征连接层,上述S802的过程可以包括:

[0176] S901、采用自适应变形层,根据各像素的卷积核、前向光流和后向光流对样本前后帧图像进行采样,得到采样结果。

[0177] 在一种可能的实施方式中,终端可以采用自适应变形层,通过一个预设的局部窗口,根据各像素的卷积核对样本前后帧图像进行采样,确定多个合成的像素,该多个合成的像素为采样结果。

[0178] 其中,终端可以根据前向光流和后向光流,确定预设的局部窗口的位置。

[0179] S902、采用自适应变形层,对样本前后帧图像、目标前向光流、目标后向光流、非局部特征进行变形处理,得到多个变形结果。

[0180] S903、采用特征连接层,对采样结果和多个变形结果进行连接,得到合成特征。

[0181] 在本申请实施例中,终端可以采用特征连接层,在通道维度上对采样结果和多个变形结果进行连接以保证特征信息不丢失,便于后续的计算和学习,可以使得生成的样本中间帧图像更加准确。

[0182] 综上所述,采用上述S901至S903的过程确定合成特征,可以使得获取的合成特征更加准确,使得合成特征所包含的局部特征和非局部特征更加丰富,便于后续对样本中间帧图像的生成。

[0183] 可选的,帧合成模块包括依次连接的多个残差模块,上述S103的过程,可以包括:

[0184] 采用多个残差模块,根据局部特征和非局部特征的合成特征,生成样本中间帧图像。

[0185] 其中,残差模块的数量可以为预设数量,例如,该预设数量可以为3。

[0186] 在一种可能的实施方式中,终端可以采用多个残差模块,根据合成特征,生成样本中间帧图像,终端还可以多个残差模块,终端线性混合了多个变形结果计算得到多个变形结果,加强了标签中间帧图像和生成中间帧图像之间的残差。

[0187] 另外,由于合成特征中所表示的多个变形结果已经经过光流对齐,这可以使得帧合成模块侧重于增强细节,使输出的中间帧图像分辨率更加高。

[0188] 在本申请实施例中,样本前帧图像可以用 I_{t-1} 表示,样本后帧图像可以用 I_{t+1} 表示,标签中间帧图像可以用 I_t 表示,样本中间帧图像可以用 \hat{I}_t 表示。

[0189] 可选的,图10为本发明实施例提供的一种模型训练方法的流程示意图,如图10所示,上述S104的过程,可以包括:

[0190] S1001、根据样本中间帧图像和对应的标签中间帧图像,确定初始神经网络模型的损失函数值。

[0191] 其中,标签中间帧图像为样本前后帧图像的标准中间帧图像。

[0192] 在一种可能的实施方式中,终端可以根据样本中间帧图像和对应的标签中间帧图像确定第一损失,根据第二损失值,继而可以对第一损失值和第二损失值进行计算,最终确定损失函数值。

[0193] S1002、根据损失函数值,对神经网络模型的参数进行调节,直至调节后的神经网络模型的损失函数值收敛。

[0194] 其中,终端可以采用预设的梯度下降算法,对神经网络模型的参数进行调节,直至调节后的神经网络模型的损失函数值收敛。

[0195] 在一些实施方式中,终端可以采用链式求导法则,对损失函数值进行计算,得到损失在神经网络模型的生成器和判别器各参数上的梯度,进而对神经网络模型中生成器和判别器的参数进行调节,以优化各参数,降低相应的损失。当调节后的神经网络模型的损失函数值收敛,得到训练好的神经网络模型。

[0196] 需要说明的是,当神经网络模型的损失函数值收敛时,可以得到训练后的神经网络模型,将变化较大的前后帧图像输入训练好的神经网络模型时,神经网络模型可以准确的生成中间帧图像。

[0197] 在本申请实施例中,终端可以根据损失函数值对生成器中的第一特征提取模块、多个第二特征提取模块、帧合成模块的参数进行优化。终端可以保存训练好的神经网络模型。

[0198] 可选的,图11为本发明实施例提供的一种模型训练方法的流程示意图,如图11所示,上述S1001的过程,可以包括:

[0199] S1101、根据样本中间帧图像和对应的标签中间帧图像,确定初始神经网络模型的像素损失和对抗损失。

[0200] 其中,像素损失用于表示样本中间帧图像和对应的标签中间帧图像之间的相似度。

[0201] 在一些实施方式中,终端可以根据样本中间帧图像和对应的标签中间帧图像,确定初始神经网络模型的对抗损失。终端可以采用预设像素损失计算公式,根据判别器的输出结果、样本中间帧图像和对应的标签中间帧图像,确定初始神经网络模型的像素损失。

[0202] 上述预设像素损失计算公式可以表示为: $\mathcal{L}_c(\Theta) = \sum_x \rho(\mathbf{I}_t(x) - \hat{\mathbf{I}}_t(x))$,其中, $\mathbf{I}_t(x)$ 为标签中间帧图像, $\hat{\mathbf{I}}_t(x)$ 为样本中间帧图像, $\rho(x) = \sqrt{x^2 + \varepsilon^2}$, x 代表图像中的坐标点, ε 是预设常数,例如该预设常数可以为 10^{-6} 。生成的中间帧图像与标签中间帧图像相似度越高, \mathcal{L}_c 越小。

[0203] S1102、根据像素损失和对抗损失,确定神经网络模型的损失函数值。

[0204] 在本发明实施例中,终端可以采用预设的加权算法,根据像素损失值和对抗损失值,确定神经网络模型的损失函数值。损失函数值可以用于对神经网络模型的参数进行优化。当损失函数值收敛时,则指示模型训练完成。

[0205] 需要说明的是,终端可以基于预设加权公式,确定损失函数值,预设加权公式可以表示为: $\mathcal{L}_G = \mathcal{L}_c + \lambda \mathcal{L}_a$,其中 λ 为权重参数, \mathcal{L}_c 为像素损失, \mathcal{L}_a 为对抗损失。权重参数越大,训练过程中对损失相关的参数的梯度越大,训练得到的神经网络模型生成的中间帧图像也随之产生变化。

[0206] 可选的,图12为本发明实施例提供的一种模型训练方法的流程示意图,如图12所示,上述S1101中,根据样本中间帧图像和对应的标签中间帧图像,确定初始神经网络模型的对抗损失的过程,可以包括:

[0207] S1201、采用判别器,确定样本中间帧图像比标签中间帧图像真实的概率,以及标签中间帧图像比样本中间帧图像虚假的概率。

[0208] 其中,判别器可以为相对判别器,相对判别器的网络结构可以为VGG-13。

[0209] 需要说明的是,标准判别器可以估计一个样本中间帧图像是真实和自然的概率,相对于标准判别器,相对判别器可以估计标签中间帧图像比样本中间帧图像更真实的概率,即判别器对样本中间帧图像和标签中间帧图像的判真概率均接近0.5,这可以增加模型训练过程的速度和稳定性。

[0210] S1202、根据真实的概率和虚假的概率,确定对抗损失。

[0211] 其中,对抗损失可以用于衡量生成器的生成能力和判别器的判断能力。终端可以采用对抗损失生成公式,根据真实的概率和虚假的概率,确定对抗损失。

[0212] 上述抗损失生成公式可以表示为: $\mathcal{L}_a(\Theta) = -\log(1 - \mathbb{D}_{Ra}(\mathbf{I}_{HR}, \mathbf{I}_{SR})) - \log(\mathbb{D}_{Ra}(\mathbf{I}_{SR}, \mathbf{I}_{HR}))$,其中 $\mathbb{D}_{Ra}(\mathbf{I}_{HR}, \mathbf{I}_{SR})$ 表示样本中间帧图像比标签中间帧图像更真实

的概率, $\mathbb{D}_{\text{Ra}}(\mathbf{I}_{\text{SR}}, \mathbf{I}_{\text{HR}})$ 表示标签中间帧图像比样本中间帧图像更虚假的概率。当 \mathcal{L}_a 收敛时,表明判别器几乎无法区分生成器生成的中间帧图像与标签中间帧图像,生成器与判别器达到均衡状态。

[0213] 综上所述,本申请实施例中生成的中间帧图像中,存在过于平滑的细节与其他不必要的视觉伪影。本发明在训练模型时,引入生成对抗网络的训练方法,增加判别器,对标签中间帧图像与样本中间帧图像进行判别,提高神经网络模型生成中间帧图像的逼真度。

[0214] 图13为本发明实施例提供的一种视频插帧方法的流程示意图,如图13所示,该视频插帧方法可以包括:

[0215] S1301、从视频中获取目标前后帧图像。

[0216] 其中,上述视频可以为不稳定帧率的视频。例如,不稳定帧率的视频可以为卡顿视频。

[0217] 在一种可能的实施方式中,终端可以采用预设算法,对视频进行检测,根据检测结果确定至少一个目标前后帧图像。终端也可以通过其他应用程序或者平台,确定视频中的目标前后帧图像,本申请实施例对此不进行具体限制。其中,目标前后帧图像包括目标前帧图像和目标后帧图像。

[0218] S1302、采用预设的神经网络模型对目标前后帧图像进行处理,得到目标中间帧图像。

[0219] 其中,神经网络模型为采用上述图1-图12任一所述的训练方法得到的神经网络模型。

[0220] 在本申请实施例中,当S1301中检测到多个目标前后帧图像时,可以将目标前后帧图像依次输入预设的神经网络模型中,以使的预设的神经网络模型可以依次对目标前后帧图像进行处理,进而依次输出各目标前后帧图像对应的目标中间帧图像。

[0221] S1303、根据目标中间帧图像对目标前后帧图像进行插帧处理。

[0222] 需要说明的是,当目标前后帧图像数量为多个时,目标中间帧图像的数量也为多个,目标中间帧图像与目标前后帧图像相对应,终端可以根据目标中间帧图像对对应的目标前后帧图像进行插帧处理,当对各目标前后帧图像均完成插帧处理时,则完成了对于视频的插帧处理。

[0223] 另外,上述获取的目标中间帧图像可以为插于目标前后帧图像之间,上述获取的目标中间帧图像该可以替换目标前后帧图像中的一个图像,终端还可以基于目标中间帧图像进行其他方式的视频插帧处理,本申请实施例对此不进行具体限制。

[0224] 综上所述,从视频中获取目标前后帧图像,采用预设的神经网络模型对目标前后帧图像进行处理,得到目标中间帧图像,根据目标中间帧图像对目标前后帧图像进行插帧处理。基于图1-图12任一所述的训练方法得到的神经网络模型生成的目标中间帧图像,对目标前后帧图像进行插帧处理,可以使得插帧后的视频更加稳定。

[0225] 可选的,上述S1301的过程,可以包括:检测视频中的至少一个重复帧标识;根据重复帧标识从视频中的多个帧图像中,确定目标前后帧图像。

[0226] 在本申请实施例中,终端可以对视屏中每个相邻两帧图像进行检测,得到检测结果,若检测结果不满足预设条件时,可以根据该相邻两帧图像的标识,确定重复帧标识,重

复帧标识可以存储于终端中。

[0227] 终端还可以存储视频中的多个帧图像,每个帧图像具有对应的帧标识,终端可以确定与重复帧标识相匹配的目标帧标识,该目标帧标识对应的帧图像为目标帧图像,该目标帧图像的前帧图像和目标帧图像的后帧图像确定为目标前后帧图像。

[0228] 需要说明的是,终端可以通过临时目录存储重复帧标识和多个帧图像。

[0229] 例如,至少一个重复帧标识也可以称为掉帧序号序列,可以表示为: $[t_1, t_2, \dots, t_M]$,终端存储的视频中的多个帧图像可以表示为: $[I_1, I_2, \dots, I_N]$,终端可以依次确定 $[t_1, t_2, \dots, t_M]$ 中对应的目标前后帧图像 $[I_{t_m-1}, I_{t_m+1}]$,目标前后帧图像对应的目标中间帧图像可以用 \hat{I}_{t_m} 表示。其中, N 为总帧数, M 为总重复帧数量, $1 \leq m \leq M \ll N$ 。

[0230] 可选的,上述检测视频中的至少一个重复帧标识,包括:

[0231] 计算视频中的相邻两个帧图像之间的像素损失;若像素损失小于或者等于预设阈值,则将后一帧图像的标识确定为重复帧标识。

[0232] 其中,相邻两个帧图像可以包括:前一帧图像和后一帧图像。

[0233] 需要说明的是,终端可以采用预设像素损失计算公式,计算视频中的相邻两个帧图像之间的像素损失。该采用预设像素损失计算公式可以表示为: $\mathcal{L}(I_t, I_{t+1}) = \sum_x \|I_t(x) - I_{t+1}(x)\|_1$,其中, $I_t(x)$ 为前一帧图像, $I_{t+1}(x)$ 为后一帧图像。

[0234] 其中,当 $\mathcal{L}(I_t, I_{t+1}) \leq T$ 时,视 I_{t+1} 为重复帧,则在记录目录中记录的重复帧标识可以为 $t+1$, $t+1$ 也可以称为掉帧序号。 T 为预设阈值参数,受图像分辨率的影响。

[0235] 可选的,上述根据目标中间帧图像对目标前后帧图像进行插帧处理的过程,可以包括:将重复帧标识对应的帧图像,替换为目标中间帧图像。

[0236] 在本申请实施例中,将所有的重复帧标识对应的帧图像,均替换为目标中间帧图像时,则可以得到的替换重复帧后的连续帧序列,继而按原始帧率合成视频,合成的视频即为帧率稳定的视频。另外,终端可以删除临时目录。

[0237] 例如,原视频中多个帧图像可以为 a, b, c, d ,对应的帧标识可以为 $1, 2, 3, 4$;重复帧图像可以为 a 和 b ,则重复帧标识可以为 2 。终端可以确定的目标前后帧图像的标识可以为 1 和 3 ,即目标前后帧图像可以为 a 和 b ,终端将 a 和 b 输入预设的神经网络模型中,得到目标中间帧图像 x ,并将 b 替换为 x ,则插帧处理后的视频中多个帧图像可以为 a, x, c, d 。

[0238] 需要说明的是,本申请实施例提供的视频插帧方法,可以部署于具有CUDA (Compute Unified Device Architecture,统一计算设备架构)核心的GPU (Graphics Processing Unit,图形处理器)终端上,也可部署于具有CUDA核心的GPU服务器端上,可以通过网络服务的方式提供调用接口。

[0239] 综上所述,本申请实施例提供的视频插帧方法,可以实现自动检测掉帧位置,可以实现不稳定帧率视频的自动优化。经过实验验证,本发明在图像质量上优于其他现有技术,且可以较好地解决不稳定帧率视频的自动优化问题。

[0240] 图14为本发明实施例提供的一种模型训练装置的结构示意图,该装置应用于神经网络模型,所述神经网络模型包括:第一特征提取模块、多个第二特征提取模块和帧合成模块,如图14所示,包括:

[0241] 提取模块1401,用于采用各第一特征提取模块,提取样本前后帧图像的局部特征;采用第二特征提取模块,提取样本前后帧图像的非局部特征;

[0242] 生成模块1402,用于采用帧合成模块,根据局部特征和非局部特征的合成特征,生成样本中间帧图像;

[0243] 训练模块1403,用于根据样本中间帧图像和对应的标签中间帧图像,对神经网络模型进行训练,得到训练后的神经网络模型。

[0244] 可选的,第二特征提取模块包括:依次连接的多组第一子模块、自注意力模块、多组第二子模块和上采样层,其中,每组第一子模块包括:降采样层和残差模块;每组第二子模块包括:上采样层和残差模块;提取模块1401,还用于采用多组第一子模块,缩小样本前后帧图像的尺寸,得到缩小后的特征图;采用自注意力模块,对缩小后的特征图进行加权,得到目标特征图,目标特征图包含有样本前后帧图像的非局部特征;采用多组第二子模块和上采样层,放大目标特征图。

[0245] 可选的,自注意力模块包括:第一卷积层、第二卷积层、第三卷积层、第四卷积层;提取模块1401,还用于分别采用第一卷积层、第二卷积层、第三卷积层对缩小后的特征图进行维度调整,得到第一调整结果、第二调整结果、第三调整结果;根据第一调整结果和第二调整结果,确定权重参数;根据权重参数和第三调整结果进行加权,得到加权结果;采用第四卷积层对加权结果进行维度恢复,得到维度恢复特征,并根据缩小后的样本前后帧图像和维度恢复特征,确定目标特征图。

[0246] 可选的,多个第一特征提取模块包括:光流估计模块、深度估计模块、核估计模块;提取模块1401,还用于采用光流估计模块,计算样本前后帧图像的前向光流和后向光流,前向光流为样本前帧图像至样本后帧图像的像素位移,后向光流为样本后帧图像至样本前帧图像的像素位移;采用深度估计模块,确定样本前后帧图像中对象间的位置关系;采用核估计模块,生成样本前后帧图像中各像素的卷积核;局部特征包括:前向光流、后向光流、对象间的位置关系、各像素的卷积核。

[0247] 可选的,神经网络模型包括深度感知投影模块,在采用帧合成模块,装置还包括:

[0248] 计算模块,用于采用深度感知投影模块,根据对象间的位置关系,重新计算前向光流和后向光流,得到目标前向光流和目标后向光流;

[0249] 确定模块,用于根据样本前后帧图像、目标前向光流、目标后向光流、非局部特征、各像素的卷积核确定合成特征。

[0250] 可选的,神经网络模型包括自适应变形层、特征连接层,确定模块,还用于采用自适应变形层,根据各像素的卷积核、前向光流和后向光流,对样本前后帧图像进行采样,得到采样结果;采用自适应变形层,对样本前后帧图像、目标前向光流、目标后向光流、非局部特征进行变形处理,得到多个变形结果;采用特征连接层,对采样结果和多个变形结果进行连接,得到合成特征。

[0251] 可选的,帧合成模块包括依次连接的多个残差模块,生成模块1402,还用于采用多个残差模块,根据局部特征和非局部特征的合成特征,生成样本中间帧图像。

[0252] 可选的,训练模块1403,还用于根据样本中间帧图像和对应的标签中间帧图像,确定初始神经网络模型的损失函数值;根据损失函数值,对神经网络模型的参数进行调节,直至调节后的神经网络模型的损失函数值收敛。

[0253] 可选的,训练模块1403,还用于根据样本中间帧图像和对应的标签中间帧图像,确定初始神经网络模型的像素损失和对抗损失,像素损失用于表示样本中间帧图像和对应的标签中间帧图像之间的相似度;

[0254] 根据像素损失和对抗损失,确定神经网络模型的损失函数值。

[0255] 可选的,训练模块1403,还用于采用判别器,确定样本中间帧图像比标签中间帧图像真实的概率,以及标签中间帧图像比样本中间帧图像虚假的概率;根据真实的概率和虚假的概率,确定对抗损失。

[0256] 图15为本发明实施例提供的一种视频插帧装置的结构示意图,如图15所示,该装置包括:

[0257] 获取模块1501,用于从视频中获取目标前后帧图像;采用预设的神经网络模型对目标前后帧图像进行处理,得到目标中间帧图像;其中,神经网络模型为采用第一方面任一的训练方法得到的神经网络模型;

[0258] 处理模块1502,用于根据目标中间帧图像对目标前后帧图像进行插帧处理。

[0259] 可选的,获取模块1501,还用于检测视频中的至少一个重复帧标识;根据重复帧标识从视频中的多个帧图像中,确定目标前后帧图像。

[0260] 可选的,获取模块1501,还用于计算视频中的相邻两个帧图像之间的像素损失,相邻两个帧图像包括:前一帧图像和后一帧图像;若像素损失小于或者等于预设阈值,则将后一帧图像的标识确定为重复帧标识。

[0261] 可选的,处理模块1502,还用于将重复帧标识对应的帧图像,替换为目标中间帧图像。

[0262] 上述装置用于执行前述实施例提供的方法,其实现原理和技术效果类似,在此不再赘述。

[0263] 以上这些模块可以是配置成实施以上方法的一个或多个集成电路,例如:一个或多个特定集成电路(Application Specific Integrated Circuit,简称ASIC),或,一个或多个微处理器(digital signal processor,简称DSP),或,一个或者多个现场可编程门阵列(Field Programmable Gate Array,简称FPGA)等。再如,当以上某个模块通过处理元件调度程序代码的形式实现时,该处理元件可以是通用处理器,例如中央处理器(Central Processing Unit,简称CPU)或其它可以调用程序代码的处理器。再如,这些模块可以集成在一起,以片上系统(system-on-a-chip,简称SOC)的形式实现。

[0264] 图16为本发明实施例提供的一种电子设备的结构示意图,如图16所示该电子设备包括:处理器1601、存储器1602。

[0265] 存储器1602用于存储程序,处理器1601调用存储器1602存储的程序,以执行上述方法实施例。具体实现方式和技术效果类似,这里不再赘述。

[0266] 可选地,本发明还提供一种程序产品,例如计算机可读存储介质,包括程序,该程序在被处理器执行时用于执行上述方法实施例。

[0267] 在本发明所提供的几个实施例中,应该理解到,所揭露的装置和方法,可以通过其它的方式实现。例如,以上所描述的装置实施例仅仅是示意性的,例如,所述单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的

相互之间的耦合或直接耦合或通信连接可以是通过一些接口,装置或单元的间接耦合或通信连接,可以是电性,机械或其它的形式。

[0268] 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0269] 另外,在本发明各个实施例中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现,也可以采用硬件加软件功能单元的形式实现。

[0270] 上述以软件功能单元的形式实现的集成的单元,可以存储在一个计算机可读取存储介质中。上述软件功能单元存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备等)或处理器(英文:processor)执行本发明各个实施例所述方法的部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(英文:Read-Only Memory,简称:ROM)、随机存取存储器(英文:Random Access Memory,简称:RAM)、磁碟或者光盘等各种可以存储程序代码的介质。

[0271] 以上仅为本发明的优选实施例而已,并不用于限制本发明,对于本领域的技术人员来说,本发明可以有各种更改和变化。凡在本发明的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。

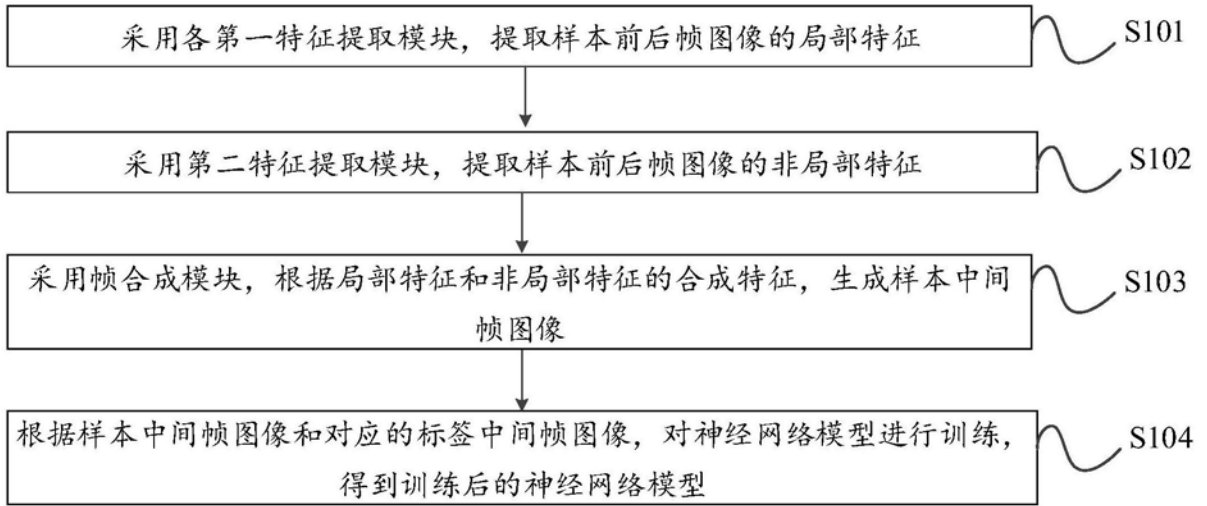


图1

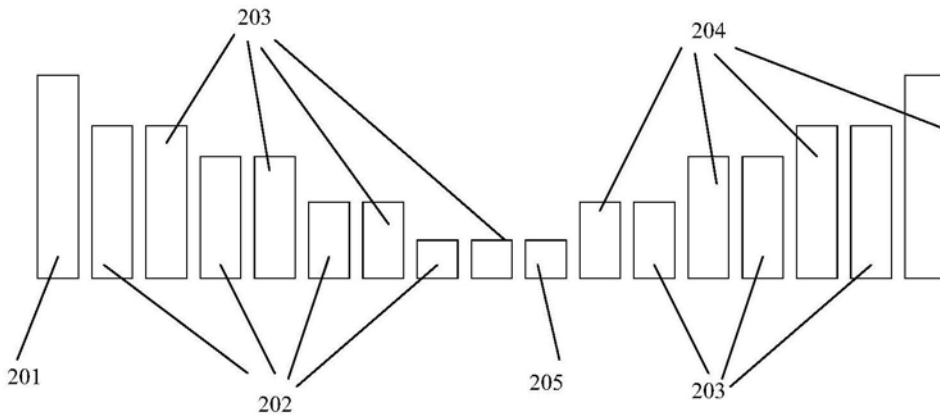


图2

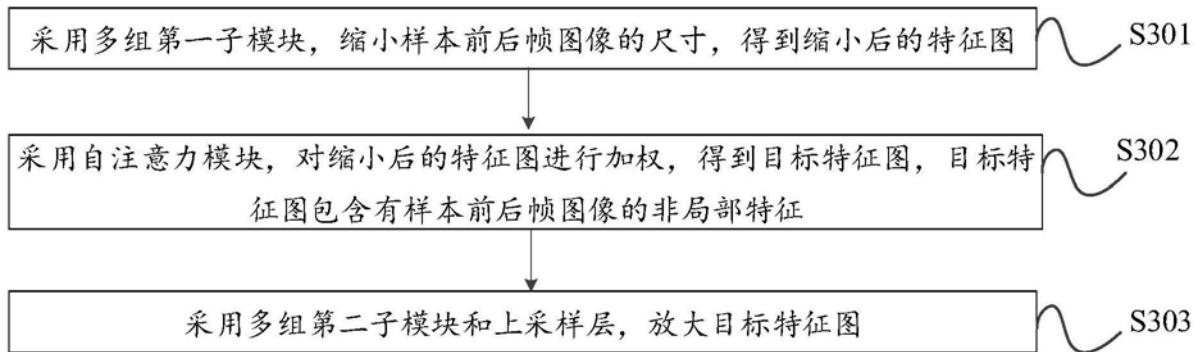


图3

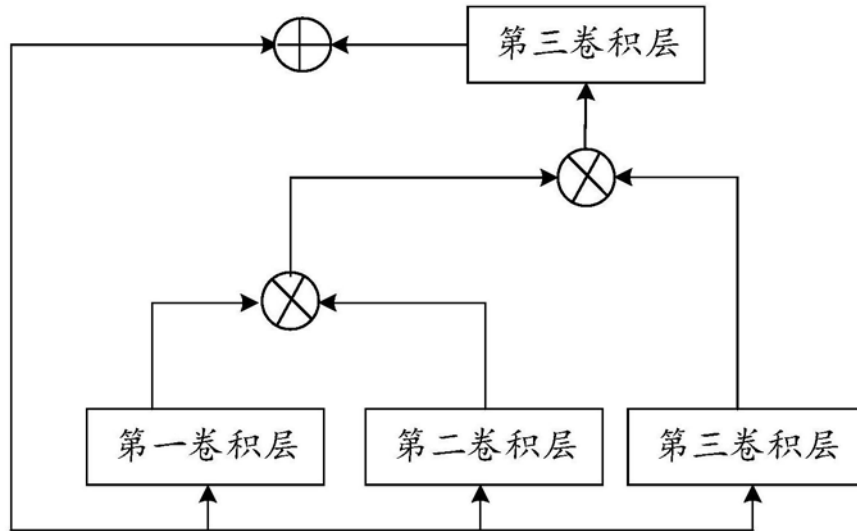


图4

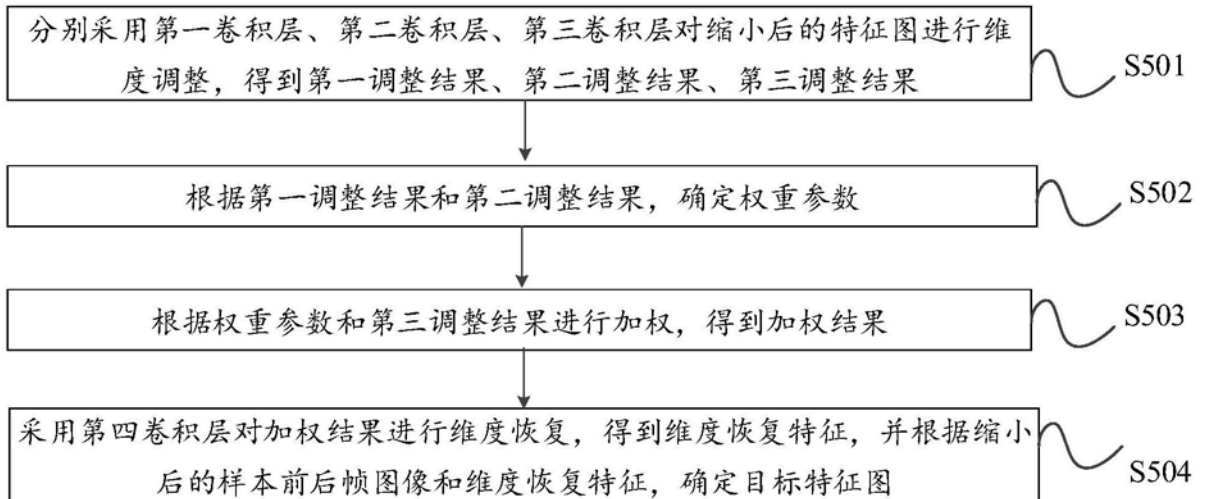


图5

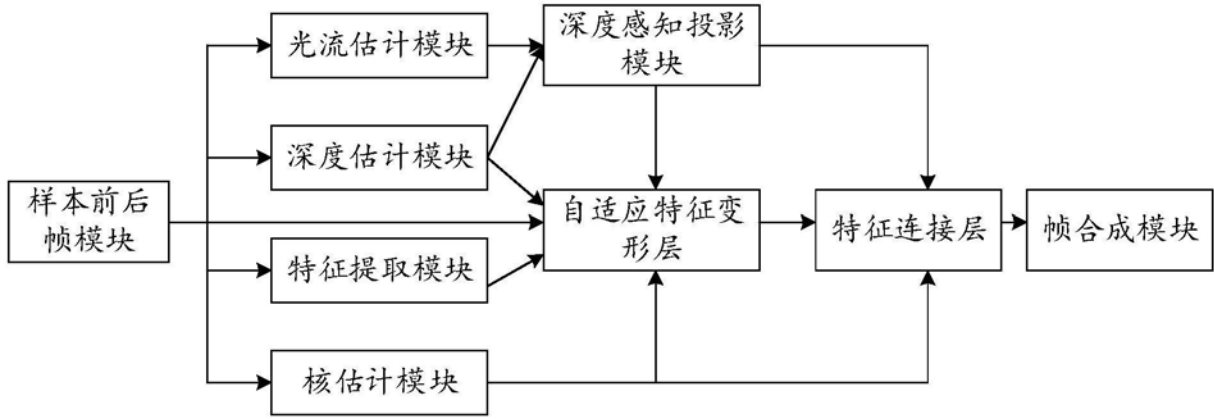


图6

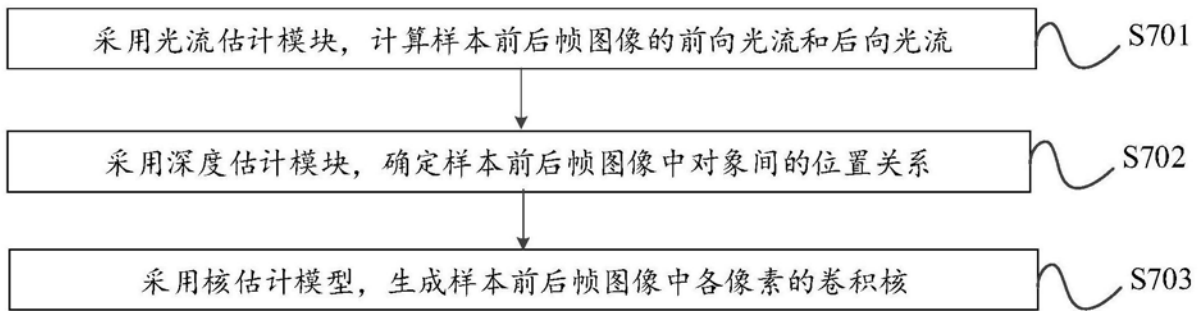


图7

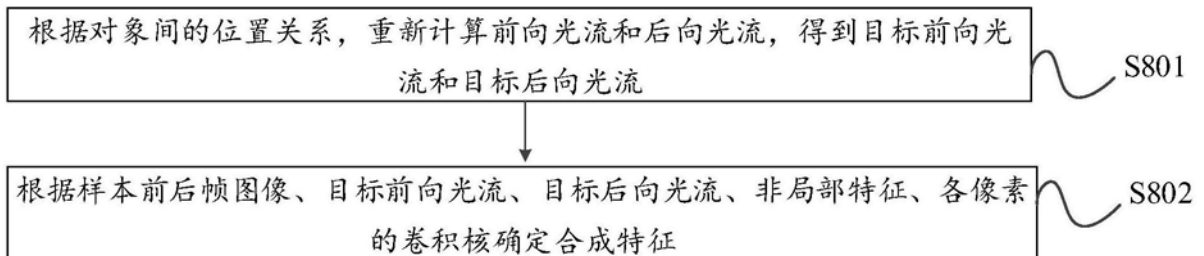


图8

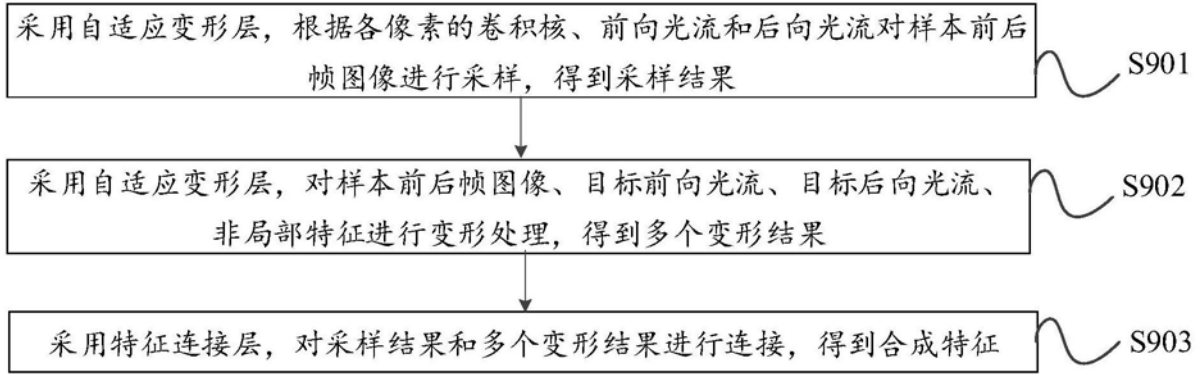


图9

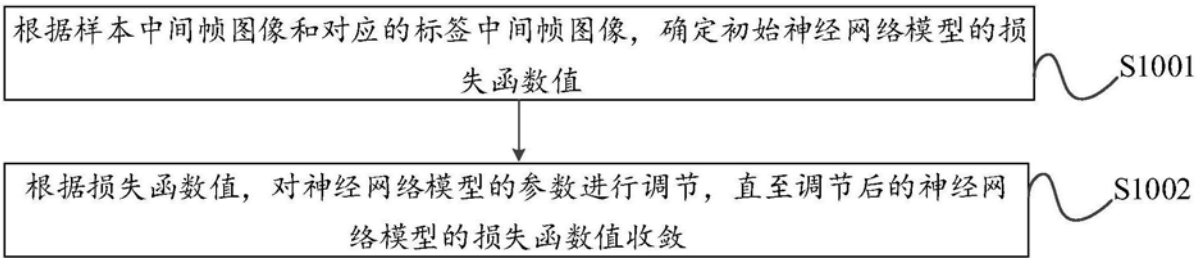


图10



图11

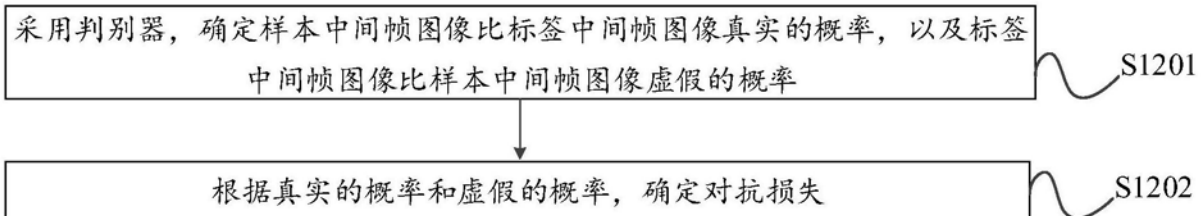


图12

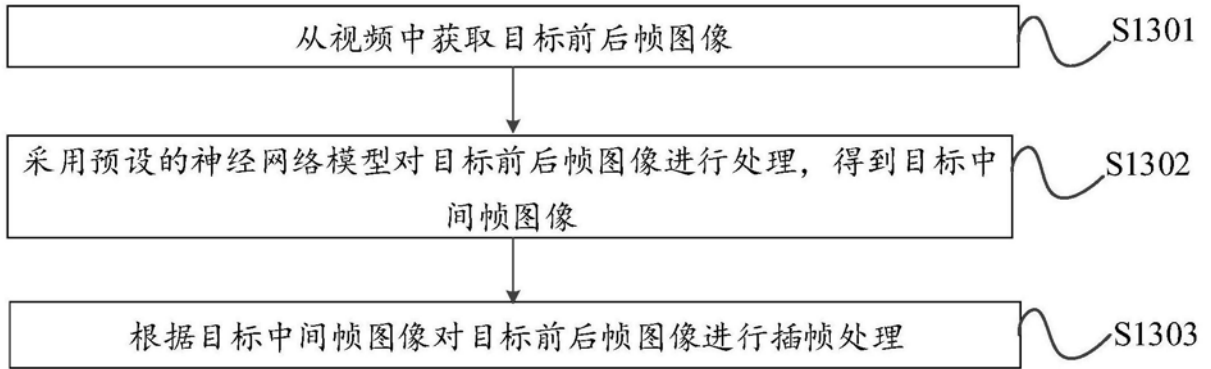


图13

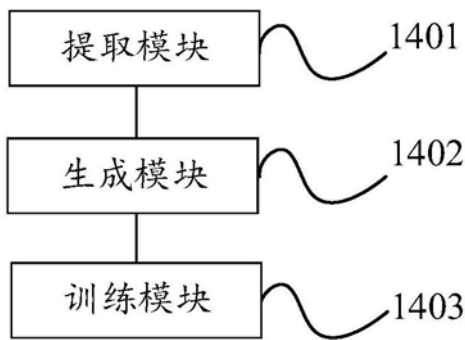


图14

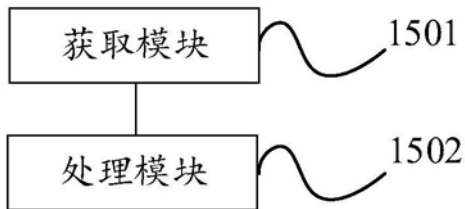


图15

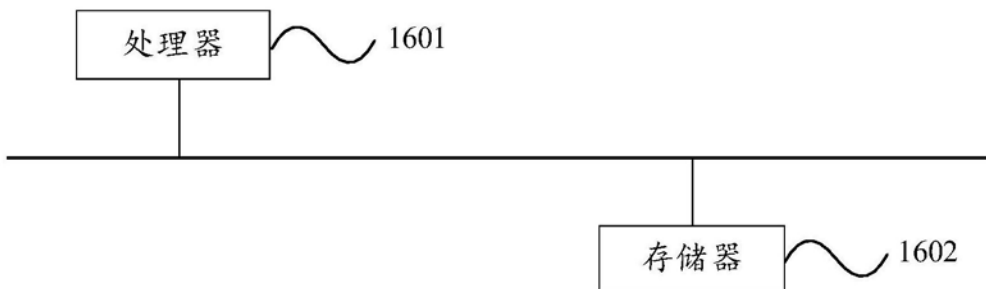


图16